TASK COMPLEXITY AND TIME PRESSURE AFFECT AIR TRAFFIC CONTROLLER'S PERFORMANCE AND WORKLOAD

Qianru Yang
Andrew R. Dattel
Dahai Liu
Embry-Riddle Aeronautical University
Daytona Beach, Florida

The purpose of this study was to investigate the effects of task complexity (TC) and time pressure (TP) on air traffic controller's (ATC) performance and mental workload. Sixteen students enrolled in an aviation college completed four scenarios which were a subset of the ATCPrep software for the AT-SAT. Fifteen performance variables were measured (e.g., conflict resolution). Additionally, NASA-TLX was used to test participants' mental workload. As expected, for most of the performance variables, high TC and high TP resulted in the lowest participant performance. For the three performance variables, TP had a differential effect on TC. Participants experienced the greatest mental workload when TC and TP were the highest. Although higher TC and higher TP was shown to have the greatest impact on performance variables, participants seemed to handle TC better than TP in several situations. When developing new technology, greater consideration should be given to TC rather than TP.

Introduction

As air traffic continues to grow, the associated demands for ATCs increase as well. ATCs play a very important role in the air traffic management system because they direct aircraft both on the ground and within the airspace. Controllers must prevent collisions, organize the flow of air traffic, and offer information to pilots. Their tasks become more cognitively demanding as traffic increases, which could compromise their performance and air traffic safety (Trapsilawati, Qu, Wickens, & Chen, 2015). Many factors have been found to have effects on ATCs' performance and mental workload, including weather, task complexity (TC), ATCs' fatigue, and time pressure (TP) (Edwards, Sharples, Wilson, & Kirwan, 2012).

ATC Task Complexity

The construct of complexity has been a largely augured matter in the ATC domain (Djokic, Lorenz, & Fricke, 2010). Many factors influencing ATCs' cognitive complexity have been indicated, such as organizational procedures, traffic environment, and display complexity (Marchitto, Di Stasi, & Cañas, 2012). ATCs play a necessary role in the safety and fluidity of the airspace as they prevent collisions, organize the flow of air traffic, and offer information to pilots. To complete their complex tasks, they use radar display to observe aircraft. When multiple aircraft show up on the display screen simultaneously, it requires greater visual attention and more working memory. This will put a high demand on mental workload for ATCs (Kaber, Perry, Segall, & Sheik-Nainar, 2007). Therefore, when all these factors are combined, it increases TC, which may influence ATCs' mental workload and their safety.

With the development of technology, aircraft and ground facilities are constantly improving and enhancing their reliability (Trapsilawati et al., 2015). However, the rate of aircraft or related equipment failure has been decreasing gradually. On the contrary, the rate of human error associated with unsafe acts has risen dramatically. (Trapsilawati et al., 2015). Mental workload assessment seems to be a recurrent problem in ATC field (Philippe, Christian, André, Sylvie, & Evelyne, 2004). Many factors can have an effect on the workload of ATCs, such as individual differences, working or living environment, TC, TP, salaries, attitude, motivation, or fatigue (Costa, 1996). ATC errors can lead to catastrophic consequences; therefore, it is important to study ATCs' performance and mental workload.

ATC Workload

An increase in the number and types of tasks is not necessarily a synonym for workload, but it also depends on individual differences, such as age, life styles, life events, work experience, and behavioral characteristics, such as mood and sleeping habits (Costa, 1996). Air traffic volume is continuing to increase worldwide. If air traffic management organizations want to meet future demands safely, they will be required to pay attention to controller's workload (Loft, Sanderson, Neal, & Mooij, 2007). Physiological measures have been used to study issues related to the effects of long-term stress on ATCs' health (Brookings, Wilson, & Swain, 1996). There are three factors that greatly affect ATC workload: time-based factors, task intensity-based factors, and operator's psychophysiological functional state (Philippe et al., 2004). High mental workload can also affect air safety due to its negative effect on ATC performance and limits traffic-handling capacities (Boag, Neal, Loft, & Halford, 2006).

Di Stasi, Marchitto, Antolí, Baccino, and Cañas (2010) found that combining different task complexities could be useful in creating different mental workload levels. The authors used an eye tracker and found that saccadic peak velocity was sensitive to variations in mental workload. During the same year, other researchers found that subjective workload hinges not only on the complexity of ATC, but also on the communication load of the ATC (Djokic et al., 2010). In addition, Fothergill and Neal (2008) used an ATC simulator and found that controllers were less likely to select the optimal solution under a high workload than under a low workload when the optimal solution was difficult to calculate. They also discovered that controllers were likely to select the optimal solution under both levels of workload when the optimal solution was easy to calculate. The following null hypotheses were tested in this study:

H_{01}: TC does not have a significant effect on an ATC's performance.
H_{02}: TP does not have a significant effect on an ATC's performance.
H_{03}: The interaction of TC and TP do not have a significant effect on an ATC's performance.
H_{04}: TC does not have a significant effect on an ATC's workload.
H_{05}: TP does not have a significant effect on an ATC's workload.
H_{06}: The interaction of TC and TP do not have a significant effect on an ATC's workload.

Method

Participants

Sixteen students at a private university in the southeastern United States were recruited. They were interested in the topic of this study. Gender and age were not factors considered. The grades (i.e., freshman, sophomore, and junior) will also not be considered for the participants.

Materials

AT-SAT software. AT-SAT is a pre-employment screening for Federal Aviation Administration ATC applicants. This software has seven cognitive tests: Air Traffic Scenarios Test, Dials Test, Analogy Test, Letter Factory Test, Angles Test, Scan Test, and Applied Mathematics Test. The Air Traffic Scenarios Test (ATST) was used in this study. In this subtest, participants should control traffic in interactive, dynamic, low-fidelity simulations of air traffic situations requiring prioritization (Dattel & King, 2010). Different scenarios can be set in the ATST. Participants handle aircraft to land at airports or go to exits efficiently. Finally, there were 15 categories scores (i.e., dependent variables), which were calculated by the software to reflect participant's performance.

NASA-TLX. In addition to the objective measures by AT-SAT of ATC's performance, their mental workload was measured by using the NASA-TLX. It is the most commonly utilized tool to measure workload (Noyes & Bruneau, 2007). The TLX is a scale with six subscales that are scored from 0 to 100. The six subscales include mental demand, physical demand, temporal demand, performance, effort, and frustration. NASA-TLX combines subscale ratings, which are weighted according to participant's subjective importance to subjects for a research (Cao, Chintamani, Pandya, & Ellis, 2009). When using NASA-TLX, participants should select two subscales of those six subscales first. These subscales are what participants find to be most relevant to the situation. Then they identify scores about these two subscales.

Procedure

Upon arrival, participants were first briefed about the purposes and procedures of the study and presented the informed consent form to review and sign. After signing the informed consent form, participants were trained how to use the AT-SAT software. The training lasted 10 minutes, which included practice trials. The test trials consisted of four scenarios, which are shown in Table 1. During the test, all aircraft in the AT-SAT software were instructed either to land at “airports” or go to “sector exits.” Low task complexity scenarios started with five aircraft; high task complexity scenarios started with 10 aircraft. In the low time pressure scenario, the airplanes were moving at a slow rate; in the high time pressure scenario, the airplanes were moving at a fast rate.

Table 1
Four Scenarios

Independent Variables	Low Task Complexity	High Task Complexity
Low Time Pressure	Low Task Complexity and Low Time Pressure	High Task Complexity and Low Time Pressure
High Time Pressure	Low Task Complexity and High Time Pressure	High Task Complexity and High Time Pressure

The study was a 2 x 2 within-subjects design. The order of four scenarios were counterbalanced using a Latin Square Design. Each scenario lasted 10 minutes. After a participant finished one scenario, he or she completed NASA-TLX, then had a 5-minute break.

Results

AT-SAT results. AT-SAT provided the following 15 performance variables. Fifteen two-way within-subject ANOVAs were run in SPSS with the alpha-level set at .05. Table 2 shows the descriptive statistics of these 15 categories scores. Table 3 shows the results of these two variables and their interaction on ATCs’ performance for 15 categories.

Table 2
Descriptive Statistics of Total Scores

	N	Minimum	Maximum	Mean	Std. Deviation
Dis Ned	64	72.10	100.00	99.45	3.54
Ti Ned	64	49.60	100.00	96.89	8.79
Conflicts	64	.00	100.00	73.26	31.88
Collisions	64	46.20	100.00	98.80	6.77
Pro Airs	64	98.40	100.00	99.83	.25
Pro Airp	64	77.80	100.00	98.82	4.36
Exit Airs	64	57.10	100.00	94.10	10.17
Exit Spd	64	57.10	100.00	96.10	8.69
Exit Alt	64	57.10	100.00	92.90	10.48
Land Des	64	.00	100.00	86.78	24.96
Land Head	64	.00	100.00	76.53	24.19
Land Spd	64	.00	100.00	81.85	23.22
Land Alt	64	.00	100.00	85.90	24.59
Set Dif	64	66.70	77.80	72.25	5.59
Total Result	64	41.60	102.60	71.64	21.85

Note. Dis Ned = Distance Needed, Ti Ned = Time Needed, Pro Airs = Prohibited Airspace Border Crossings, Pro Airp = Prohibited Airport Border Crossings, Exit Airs = Exiting the Airspace Correct Destination, Exit Spd = Exiting the Correct Speed, Exit Alt = Exiting the Correct Altitude, Land Des = Landing at Airports Correct Destination, Land Head = Landing at Airports Correct Headings, Land Spd = Landing at Airports Speed, Land Alt = Landing at Airports Correct Altitude, Set Dif = Set up Difficulty.

Table 3
Significance of Independent Variables and Performance

	Task Complexity			Time Pressure			Task Complexity and Time Pressure		
Per Va	*F*	*p*	Rejected or Retained	*F*	*p*	Rejected or Retained	*F*	*p*	Rejected or Retained
Dis Ned	F (1, 15) = .70,	.42	Retained	F (1, 15) = .70	.42	Retained	F (1, 15) = 1.59	.23	Retained
Ti Ned	F (1, 15) = .08	.78	Retained	F (1, 15) = 4.99	< .05	Rejected*	F (1, 15) = .89	.36	Retained
Conflicts	F (1, 15) = 49.36	< .05	Rejected*	F (1, 15) = 92.78	< .05	Rejected*	F (1, 15) = 27.09	< .05	Rejected*
Collisions	F (1, 15) = 1.42	.25	Retained	F (1, 15) = 1.65	.22	Retained	F (1, 15) = 1.64	.22	Retained
Pro Airs	F (1, 15) = .002	.97	Retained	F (1, 15) = 1.31	.27	Retained	F (1, 15) = 2.74	.12	Retained
Pro Airp	F (1, 15) = 1.47	.25	Retained	F (1, 15) = 1.48	.24	Retained	F (1, 15) = .13	.73	Retained
Exit Airs	F (1, 15) = 9.58	< .05	Rejected*	F (1, 15) = 16.59	< .05	Rejected*	F (1, 15) = 36.48	< .05	Rejected*
Exit Spd	F (1, 15) = 25.54	< .05	Rejected*	F (1, 15) = 30.67	< .05	Rejected*	F (1, 15) = 25.54	< .05	Rejected*
Exit Alt	F (1, 15) = 6.67	< .05	Rejected*	F (1, 15) = 14.12	< .05	Rejected*	F (1, 15) = 19.14	< .05	Rejected*
Land Des	F (1, 15) = 22.75	< .05	Rejected*	F (1, 15) = 20.55	< .05	Rejected*	F (1, 15) = 18.03	< .05	Rejected*
Land Head	F (1, 15) = 6.85	< .05	Rejected*	F (1, 15) = 3.09	.10	Retained	F (1, 15) = 1.27	.28	Retained
Land Spd	F (1, 15) = 8.31	< .05	Rejected*	F (1, 15) = 3.79	.07	Retained	F (1, 15) = 6.88	< .05	Rejected*
Land Alt	F (1, 15) = 18.10	< .05	Rejected*	F (1, 15) = 15.48	< .05	Rejected*	F (1, 15) = 13.86	< .05	Rejected*
Set Dif	F (1, 15) = 1	.33	Retained	F (1, 15) = 225	< .05	Rejected*			
To Re	F (1, 15) = 179.23	< .05	Rejected*	F (1, 15) = 2.31	.15	Retained	F (1, 15) = 38.43	< .05	Rejected*

Note. Per Va = Performance Variables, Dis Ned = Distance Needed, Ti Ned = Time Needed, Pro Airs = Prohibited Airspace Border Crossings, Pro Airp = Prohibited Airport Border Crossings, Exit Airs = Exiting the Airspace Correct Destination, Exit Spd = Exiting the Correct Speed, Exit Alt = Exiting the Correct Altitude, Land Des = Landing at Airports Correct Destination, Land Head = Landing at Airports Correct Headings, Land Spd = Landing at Airports Speed, Land Alt = Landing at Airports Correct Altitude, Set Dif = Set up Difficulty, To Re = Total Result.

NASA-TLX results. There are six subscales in NASA-TLX. The NASA-TLX provides two results for participants' mental workload. One of the six subscales were the most relevant to workload. The other one of the results was mean value of overall workload. Table 4 shows the description of overall workload.

Table 4
Description of Overall Workload

	Valid	Missing	Mean	Median	SD	Min	Max
LT_LTP	16	0	89.13	85	49.86	30	175
LT_HTP	16	0	133.31	147.50	43.02	60	188
HT_LTP	16	0	120.88	132.50	46.29	30	175
HT_HTP	16	0	164.81	172.50	33.89	95	200

Note. SD= Std. Deviation, Min = Minimum, Max = Maximum, LT_LTP = Low Task Complexity and Low Time Pressure, LT_HTP = Low Task Complexity and High Time Pressure, HT_LTP = High Task Complexity and Low Time Pressure, HT_HTP = High Task Complexity and High Time Pressure.

A two-way within-subject ANOVA was run in SPSS with the alpha level set at .05. The results that were analyzed the overall score. Therefore, the results of overall workload were: F (1, 15) = 14.72, p < .05 for TC; F (1, 15) = 45.86, p < .05 for TP.

Discussion and Conclusion

Task complexity and time pressure affect performance.

There are 15 categories for AT-SAT to reflect ATC performance. The data yielded some intriguing findings. Results show that distance, airspace border, and airport border were not affected by TC and TP. Second, when TC is higher and TP is lower, ATCs had better results of Exiting the Airspace Correct Destination than when TP is higher; therefore, the level of TP had a greater negative impact on performance than TC. For the "Total Result" variable, when TP is lower and TC at the same level, ATCs performed better. In addition, when TC was higher, regardless of TP level, ATCs had better performance. This indicates that an increased number of aircraft yields greater workload, high TC may promote better performance.

TC and TP affect workload.

Results showed that mental workload occurred more frequently than any of the other five subscales (i.e., physical demand, temporal demand, performance, effort, and frustration). This means that mental workload is the most relevant subscale of ATCs workload. As expected, high TC or high TP had a greater effect on ATCs' workload than low TC or TP respectively.

Future research should consider these suggestions for improvement. First, conducting this experiment utilizing trained ATCs may yield more reliable results. Second, scenarios may have different levels of difficulty, such as low, medium, and high. Moreover, higher TC is not necessarily bad in all situations. Therefore, when doing selecting and training of ATC in the future, TC might be considered more than TP.

Reference

Boag, C., Neal, A., Loft, S., & Halford, G. S. (2006). An analysis of relational complexity in an air traffic control conflict detection task. *Ergonomics*, *49*, 1508-1526. doi: 10.1080/00140130600779744

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology, 42*, 361-377. doi: 10.1016/0301-0511(95)05167-8

Cao, A., Chintamani, K. K., Pandya, A. K., & Ellis, R. D. (2009). NASA TLX: Software for assessing subjective mental workload. *Behavior Research Methods, 41*, 113-117. doi: 10.3758/BRM.41.1.113

Costa, G. (1996). Occupational stress and stress prevention in air traffic control. International Labour Organization.

Dattel, A. R., & King, R. E. (2010). The impact of reweighting the AT-SAT subtests on group score differences. *The International Journal of Aviation Psychology*, *20*, 111-123. doi: 10.1080/10508411003617748

Djokic, J., Lorenz, B., & Fricke, H. (2010). Air traffic control complexity as workload driver. *Transportation Research Part C: Emerging Technologies, 18*, 930-936. doi: 10.1016/j.trc.2010.03.005

Di Stasi, L. L., Marchitto, M., Antolí, A., Baccino, T., & Cañas, J. J. (2010). Approximation of on-line mental workload index in ATC simulated multitasks. *Journal of Air Transport Management*, *16*, 330-333. doi: 10.1016/j.jairtraman.2010.02.004

Edwards, T., Sharples, S., Wilson, J. R., & Kirwan, B. (2012). Factor interaction influences on human performance in air traffic control: The need for a multifactorial model. *Work*, *41*, 159-166. doi: 10.3233/WOR-2012-0151-159

Fothergill, S., & Neal, A. (2008). The effect of workload on conflict decision making strategies in air traffic control. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 52*, 39-43. doi:10.1177/154193120805200110

Kaber, D. B., Perry, C. M., Segall, N., & Sheik-Nainar, M. A. (2007). Workload state classification with automation during simulated air traffic control. *The International Journal of Aviation Psychology, 17*, 371-390. doi: 10.1080/10508410701527860

Loft, S., Sanderson, P., Neal, A., & Mooij, M. (2007). Modeling and predicting mental workload in en route air traffic control: Critical review and broader implications. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 49*, 376-399. doi: 10.1518/001872007X197017

Marchitto, M., Di Stasi, L. L., & Cañas, J. J. (2012). Ocular movements under taskload manipulations: Influence of geometry on saccades in air traffic control simulated tasks. *Human Factors and Ergonomics in Manufacturing & Service Industries, 22,* 407-419. doi: 10.1002/hfm.20320

Noyes, J. M., & Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics, 50,* 514-519. doi: 10.1080/00140130701235232

Philippe, A., Christian, C., André, D., Sylvie, A., & Evelyne, V. (2004). Mental workload in air traffic control: An index constructed from field tests. *Aviation, Space, and Environmental Medicine, 75*, 333-341.

Trapsilawati, F., Qu, X., Wickens, C. D., & Chen, C. (2015). Human factors assessment of conflict resolution aid reliability and time pressure in future air traffic control. *Ergonomics, 58,* 897-908. doi: 10.1080/00140139.2014.997301